On the Impact of a Small Initial Population Size in the IPOP Active CMA-ES with Mirrored Mutations on the Noiseless BBOB Testbed

Dimo Brockhoff, Anne Auger, Nikolaus Hansen

On the Impact of a Small Initial Population Size in the IPOP Active CMA-ES with Mirrored Mutations on the Noiseless BBOB Testbed

Dimo Brockhoff
INRIA Lille - Nord Europe
Dolphin team
59650 Villeneuve d'Ascq
France
dimo.brockhoff@inria.fr

Anne Auger
Projet TAO, INRIA
Saclay—Ile-de-France
LRI, Bât 490, Univ. Paris-Sud
91405 Orsay Cedex, France
anne.auger@inria.fr

Nikolaus Hansen
Projet TAO, INRIA
Saclay—Ile-de-France
LRI, Bât 490, Univ. Paris-Sud
91405 Orsay Cedex, France
nikolaus.hansen@inria.fr

ABSTRACT

Active Covariance Matrix Adaptation and Mirrored Mutations have been independently proposed as improved variants of the well-known optimization algorithm Covariance Matrix Adaptation Evolution Strategy (CMA-ES) for numerical optimization. This paper investigates the impact of the algorithm's population size when both active covariance matrix adaptation and mirrored mutation are used in the CMA-ES. To this end, we compare the CMA-ES with standard population size λ, i.e., $\lambda = 4 + \lfloor 3 \log(D) \rfloor$ with a version with half this population size where D is the problem dimension.

Categories and Subject Descriptors

G.1.6 [**Numerical Analysis**]: Optimization—*global optimization, unconstrained optimization*; F.2.1 [**Analysis of Algorithms and Problem Complexity**]: Numerical Algorithms and Problems

General Terms

Algorithms

Keywords

Benchmarking, Black-box optimization

1. INTRODUCTION

The IPOP-CMA-ES [2] has the special feature of increasing the population size of the CMA-ES algorithm at each restart. Together with a standard population size of $\lambda^s = 4 + \lfloor 3 \log(D) \rfloor$ where D is the problem dimension, the IPOP-CMA-ES is a (nearly) parameterless algorithm that automatically restarts CMA-ES with increased population size if the given size is not sufficient to solve the problem at hand.

More recently, an active covariance matrix adaptation update has been proposed for CMA-ES [9] and mirrored mutations with pairwise selection and selective mirroring have been suggested for evolution strategies with weighted recombination [1]. While the former one allows for negative weights in the covariance matrix update for bad mutations, the latter mirrors the bad mutations and evaluates them again before to proceed.

The combination of both approaches into the IPOP-CMA-ES with active covariance matrix adaptation and mirrored mutations, denoted by CMA_{ma}, has been introduced and tested empirically in an accompanying paper [3]. Here, we test how a different starting population size influences the performance of this algorithm. A previous study showed that in the $(1, \lambda)$-ES, the largest effect of mirrored mutations is observed for small population sizes, i.e. $\lambda = 2$ and $\lambda = 4$ [4]. Hence, we could conjecture that in the IPOP-CMA-ES with mirrored mutations, a positive effect on the performance can be observed if the initial population size is chosen smaller than the standard size of λ^s. To test this hypothesis, we run the CMA_{ma} with an initial population size of λ^s and compare it with the CMA_{mah} that employs an initial population size of $\lfloor \lambda^s/2 \rfloor$ on the noiseless BBOB test bed [6].

The algorithms are described in more detail in Sec. 2. Section 3 gives the mandatory results of the BBOB timing experiments while Sec. 4 presents the general results of the comparison. Section 5 concludes the paper.

2. TESTED CMA-ES VARIANTS

We tested two variants of the IPOP-CMA-ES with active covariance matrix adaptation and mirrored mutations: the CMA_{ma} with standard initial population size $\lambda^s = 4 + \lfloor 3 \log(D) \rfloor$ and the CMA_{mah} with reduced initial population size $\lfloor \lambda^s/2 \rfloor$. Both implementations can be downloaded from `http://canadafrance.gforge.inria.fr/mirroring/` in the used version 3.54.beta.mirrors. Besides the difference in the initial population size, the number of restarts is increased to 10 for the CMA_{mah} instead of 9 for the CMA_{ma} to allow the final restarts of both algorithms to operate with the same (range of) population size. All other parameters are equal for the two algorithms and, besides $2 \cdot 10^5 \cdot D$ as the maximal number of function evaluations, chosen according to the standard recommendations for the CMA-ES. For more details of the algorithm, see also the accompanying paper [3].

3. TIMING EXPERIMENTS

In order to see the dependency of the algorithms on the problem dimension, the requested BBOB'2012 timing experiment has been performed for the two algorithms CMA_{ma} and CMA_{mah} on an Intel Core2 Duo T9600 laptop with 2.80GHz, 4.0GB of RAM, and MATLAB R2008b on Windows Vista SP2. The algorithms have been restarted for up to $2 \cdot 10^5 D$ function evaluations until 30 seconds have passed. The per-function-evaluation-runtimes were 16; 16; 11; 6.5; 4.2; 4.6; and 7.2 times 10^{-4} seconds for the CMA_{ma} and 21; 19; 11; 8.3; 6.1; 5.7 and 11 times 10^{-4} seconds for the CMA_{mah} in 2, 3, 5, 10, 20, 40, and 80 dimensions respectively.

4. RESULTS

Results from experiments according to [6] on the benchmark functions given in [5, 7] are presented in Figures 1, 2 and 3 and in Tables 1. The **expected running time (ERT)**, used in the figures and table, depends on a given target function value, $f_t = f_{\text{opt}} + \Delta f$, and is computed over all relevant trials as the number of function evaluations executed during each trial while the best function value did not reach f_t, summed over all trials and divided by the number of trials that actually reached f_t [6, 10]. **Statistical significance** is tested with the rank-sum test for a given target Δf_t (10^{-8} as in Figure 1) using, for each trial, either the number of needed function evaluations to reach Δf_t (inverted and multiplied by -1), or, if the target was not reached, the best Δf-value achieved, measured only up to the smallest number of overall function evaluations for any unsuccessful trial under consideration.

The first observation is the fact that both algorithm variants behave quite similar with only a few cases where the differences are statistically significant. The two main exceptions are the sphere function (f_1) for which the variant with smaller initial population size is about 25% faster in all dimensions and all difficult targets and the discus function (f_{11}) where the variant with standard population size is about 20% faster in 5D and 10% faster in 20D (see Table 1). Figure 1 reveals a few more statistically significant differences for the target value 10^{-8}: while the algorithm with standard population size is faster for several lower dimensions (f_2 in 2D and 3D, f_{10} in 2D, 3D, and 10D, f_{13} in 2D, 3D, and 5D, f_{17} in 3D and 10D, f_{18} in 3D and 20D) as well as on f_6 in 20D and 40D and for f_{14} in all dimensions but 20, the algorithm with reduced initial population size is sometimes faster for larger dimensions (f_5 in 20D and 40D, f_8 and f_{12} in 40D). Furthermore, one can observe that, in 20D, unsuccessful runs occur for eight of the 24 functions and the functions f_3, f_4, and f_{19}–f_{24} cannot be solved by both algorithms in any of the 15 runs. When compared to the best algorithm of the BBOB'2009 exercise, both algorithms significanly improve the performance on f_{10} (faster by a factor of 1.4), f_{14} (factor of ≥ 1.5), and f_{11} and f_{15} (factor of > 2, all results in 40D) which is mainly due to the active covariance matrix adaptation [8].

5. CONCLUSIONS

When investigating the impact of the initial population size in the IPOP-CMA-ES with active covariance matrix adaptation and mirrored mutation, no general recommendation towards one of the two algorithms CMA_{ma} and CMA_{mah} can be made. While a lower population size is generally helpful on the sphere function and less effective on the discus function, the positive effect of the lower population size is often more pronounced for larger dimensions with the exception of the attractive sector function where the opposite is the case. As a general conclusion, we remark that the change of the initial population size has overall comparatively small effects.

6. REFERENCES

[1] A. Auger, D. Brockhoff, and N. Hansen. Mirrored Sampling in Evolution Strategies With Weighted Recombination. In *Genetic and Evolutionary Computation Conference (GECCO 2011)*, pages 861–868. ACM, 2011.

[2] A. Auger and N. Hansen. A Restart CMA Evolution Strategy With Increasing Population Size. In *Congress on Evolutionary Computation (CEC 2005)*, volume 2, pages 1769–1776. IEEE Press, 2005.

[3] D. Brockhoff, A. Auger, and N. Hansen. On the Effect of Mirroring in the IPOP Active CMA-ES on the Noiseless BBOB Testbed. In *GECCO (Companion) workshop on Black-Box Optimization Benchmarking (BBOB'2012)*. ACM, 2012.

[4] D. Brockhoff, A. Auger, N. Hansen, D. V. Arnold, and T. Hohm. Mirrored Sampling and Sequential Selection for Evolution Strategies. In *Conference on Parallel Problem Solving from Nature (PPSN XI)*, pages 11–21. Springer, 2010.

[5] S. Finck, N. Hansen, R. Ros, and A. Auger. Real-parameter black-box optimization benchmarking 2009: Presentation of the noiseless functions. Technical Report 2009/20, Research Center PPE, 2009. Updated February 2010.

[6] N. Hansen, A. Auger, S. Finck, and R. Ros. Real-parameter black-box optimization benchmarking 2012: Experimental setup. Technical report, INRIA, 2012.

[7] N. Hansen, S. Finck, R. Ros, and A. Auger. Real-parameter black-box optimization benchmarking 2009: Noiseless functions definitions. Technical Report RR-6829, INRIA, 2009. Updated February 2010.

[8] N. Hansen and R. Ros. Benchmarking a weighted negative covariance matrix update on the BBOB-2010 noiseless testbed. In *Genetic and Evolutionary Computation Conference (GECCO 2010)*, pages 1673–1680, New York, NY, USA, 2010. ACM.

[9] G. Jastrebski and D. Arnold. Improving evolution strategies through active covariance matrix adaptation. In *IEEE Congress on Evolutionary Computation (CEC 2006)*, pages 2814–2821, 2006.

[10] K. Price. Differential evolution vs. the functions of the second. In *Proceedings of the IEEE International Congress on Evolutionary Computation (ICEO)*, pages 153–157, 1997.

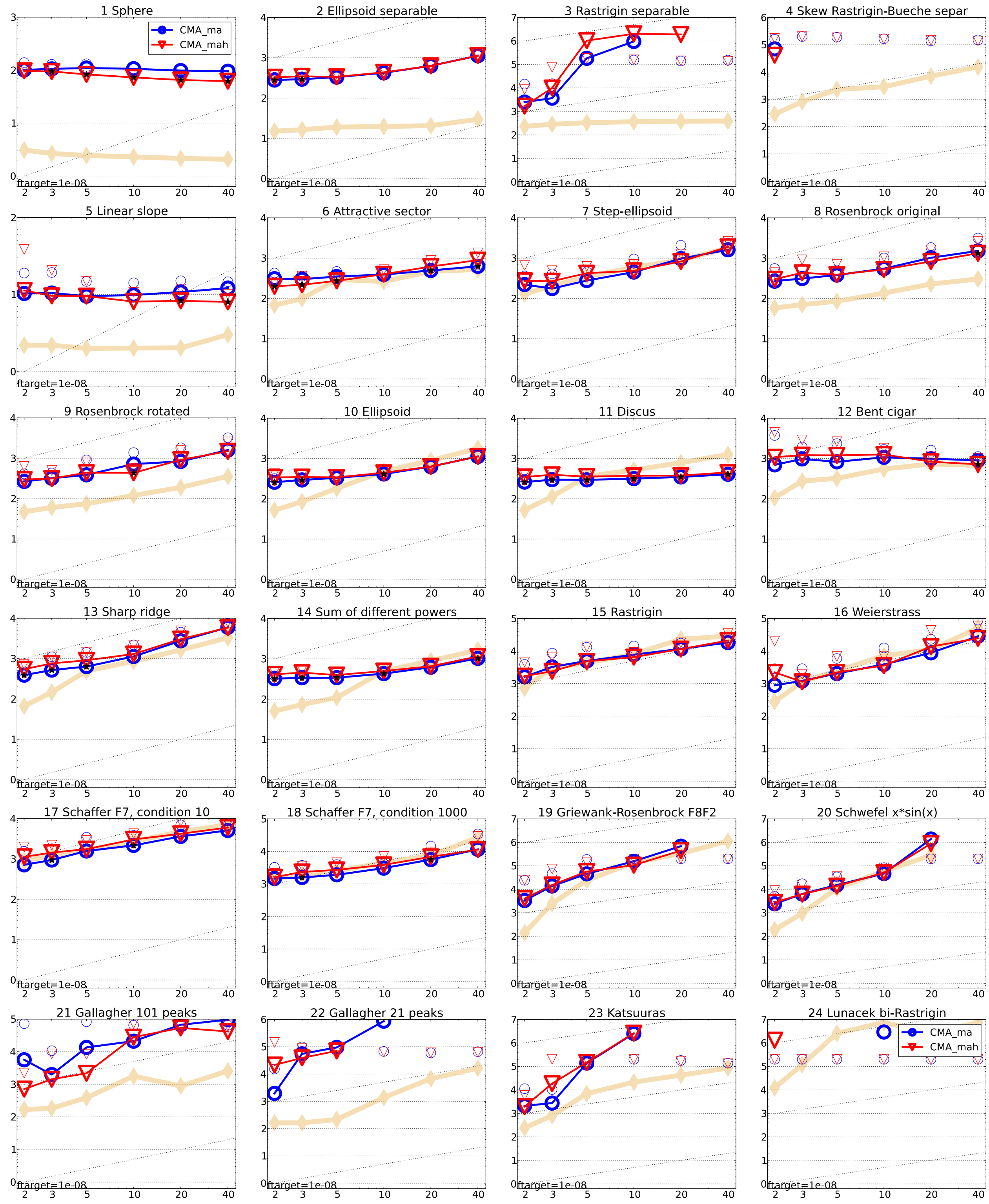

Figure 1: Expected running time (ERT in number of f-evaluations) divided by dimension for target function value 10^{-8} as $\log_{10}$ values versus dimension. Different symbols correspond to different algorithms given in the legend of f_1 and f_{24}. Light symbols give the maximum number of function evaluations from the longest trial divided by dimension. Horizontal lines give linear scaling, slanted dotted lines give quadratic scaling. Black stars indicate statistically better result compared to all other algorithms with $p < 0.01$ and Bonferroni correction number of dimensions (six). Legend: $\circ$:CMA$_{\mathrm{ma}}$, $\triangledown$:CMA$_{\mathrm{mah}}$.

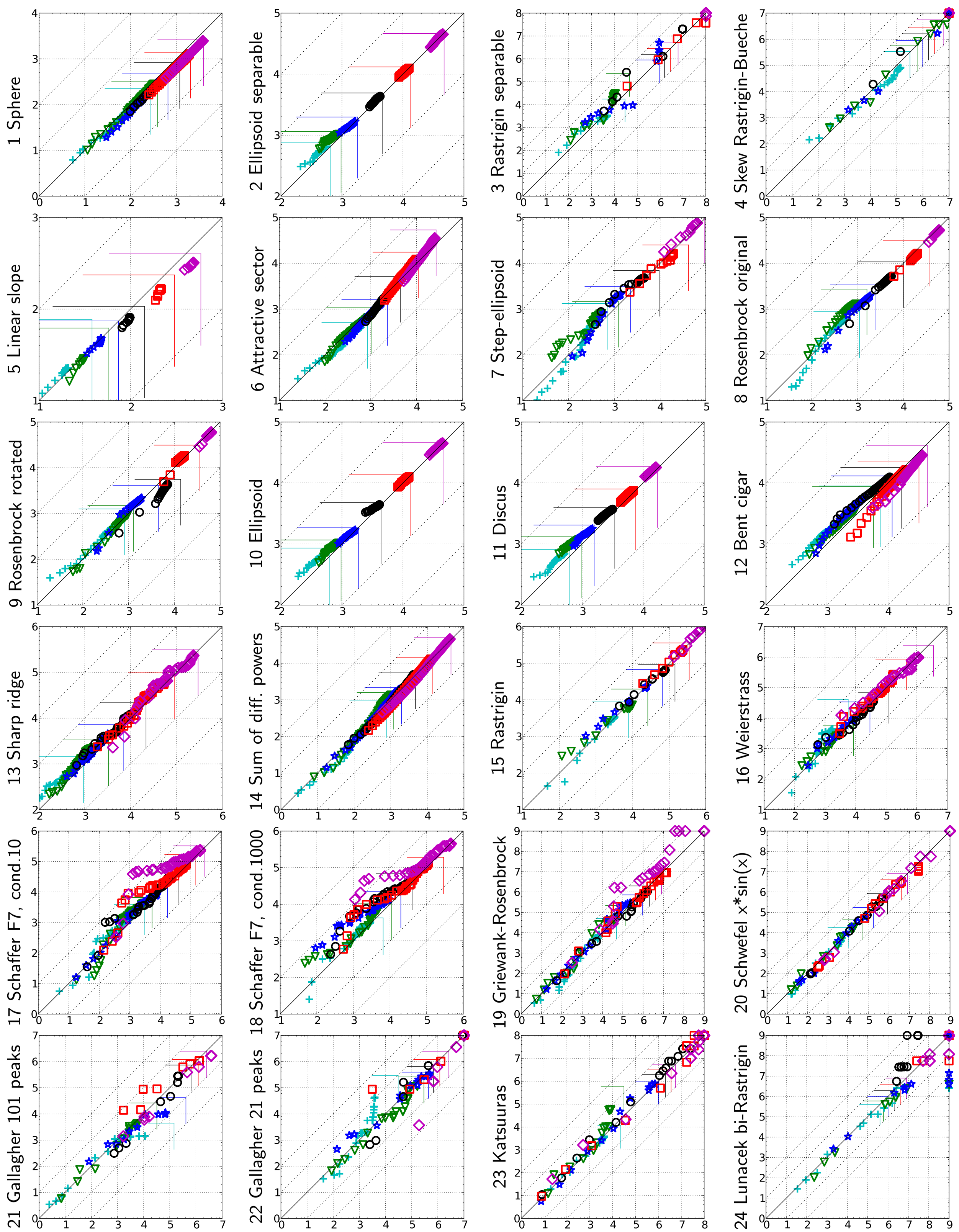

Figure 2: Expected running time (ERT in $\log_{10}$ of number of function evaluations) of CMA$_{\text{ma}}$ (x-axis) versus CMA$_{\text{mah}}$ (y-axis) for 46 target values $\Delta f \in [10^{-8}, 10]$ in each dimension on functions f_1–f_{24}. Markers on the upper or right edge indicate that the target value was never reached. Markers represent dimension: 2:+, 3:▽, 5:⋆, 10:○, 20:□, 40:◇.

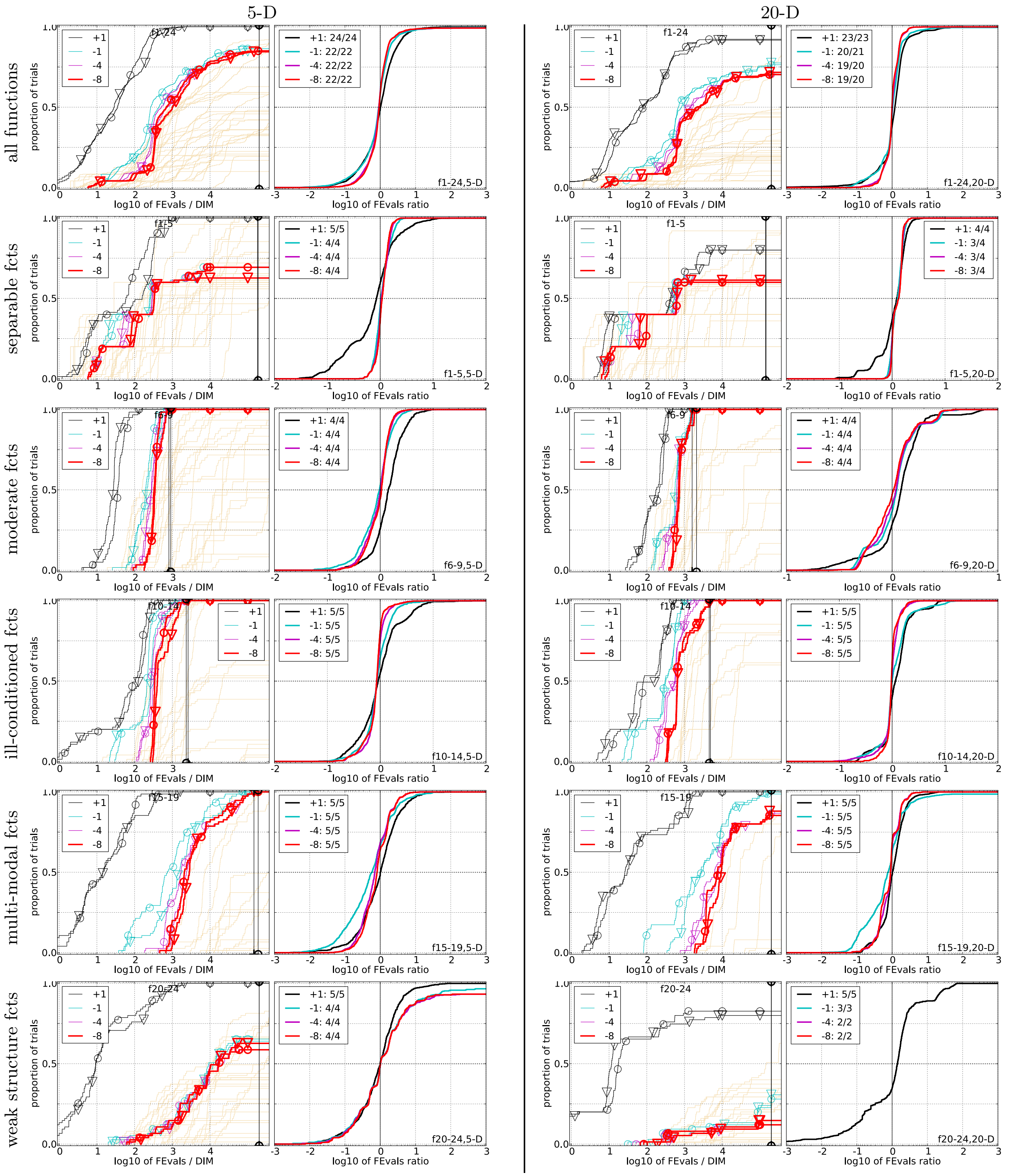

Figure 3: Empirical cumulative distributions (ECDF) of run lengths and speed-up ratios in 5-D (left) and 20-D (right). Left sub-columns: ECDF of the number of function evaluations divided by dimension D (FEvals/D) to reach a target value $f_{\rm opt} + \Delta f$ with $\Delta f = 10^k$, where $k \in \{1, -1, -4, -8\}$ is given by the first value in the legend, for CMA$_{\rm ma}$ ($\circ$) and CMA$_{\rm mah}$ ($\triangledown$). Light beige lines show the ECDF of FEvals for target value $\Delta f = 10^{-8}$ of all algorithms benchmarked during BBOB-2009. Right sub-columns: ECDF of FEval ratios of CMA$_{\rm ma}$ divided by CMA$_{\rm mah}$, all trial pairs for each function. Pairs where both trials failed are disregarded, pairs where one trial failed are visible in the limits being > 0 or < 1. The legends indicate the number of functions that were solved in at least one trial (CMA$_{\rm ma}$ first).

5-D

Δf	1e+1	1e-1	1e-3	1e-5	1e-7	#succ
$\mathbf{f_1}$	11	12	12	12	12	15/15
1: CMA	2.7(3)	11(4)	22(4)	31(5)	40(7)	15/15
2: CMA	1.7(1)	8.1(2)	**16**(2)$^{\star 2}$	**23**(2)$^{\star 3}$	**30**(3)$^{\star 3}$	15/15
$\mathbf{f_2}$	83	88	90	92	94	15/15
1: CMA	11(3)	14(2)	15(2)	16(2)	17(1)	15/15
2: CMA	13(3)	15(2)	16(2)	17(2)	17(2)	15/15
$\mathbf{f_3}$	716	1637	1646	1650	1654	15/15
1: CMA	**0.74**(1)*	556(737)	554(725)	553(770)	552(728)	7/15
2: CMA	2.4(2)	3145(3743)	3129(3709)	3121(3761)	3114(3650)	2/15
$\mathbf{f_4}$	809	1688	1817	1886	1903	15/15
1: CMA	1.7(2)	∞	∞	∞	∞ *8.8e5*	0/15
2: CMA	2.5(2)	∞	∞	∞	∞ *8.7e5*	0/15
$\mathbf{f_5}$	10	10	10	10	10	15/15
1: CMA	3.4(0.8)	4.8(2)	4.8(2)	4.8(2)	4.8(2)	15/15
2: CMA	3.3(2)	4.7(2)	4.8(2)	4.8(2)	4.8(2)	15/15
$\mathbf{f_6}$	114	281	580	1038	1332	15/15
1: CMA	2.4(1)	2.2(0.9)	1.6(0.4)	1.2(0.2)	1.2(0.2)	15/15
2: CMA	1.7(1.0)	1.7(0.5)	1.3(0.4)	0.94(0.3)	**0.92**(0.2)*	15/15
$\mathbf{f_7}$	24	1171	1572	1572	1597	15/15
1: CMA	5.1(3)	0.84(0.6)	0.76(0.5)	0.76(0.5)	0.82(0.5)	15/15
2: CMA	3.9(3)	1.4(0.7)	1.2(0.6)	1.2(0.6)	1.2(0.6)	15/15
$\mathbf{f_8}$	73	336	391	410	422	15/15
1: CMA	2.7(1)	3.7(2)	4.0(1)	4.2(1)	4.4(1)	15/15
2: CMA	1.8(0.6)	4.0(3)	4.3(2)	4.4(2)	4.5(2)	15/15
$\mathbf{f_9}$	35	214	300	335	369	15/15
1: CMA	5.9(2)	6.0(2)	5.4(2)	5.3(1)	5.1(1)	15/15
2: CMA	4.4(1)	7.5(5)	6.3(3)	6.1(3)	5.8(3)	15/15
$\mathbf{f_{10}}$	349	574	626	829	880	15/15
1: CMA	2.5(0.8)	2.1(0.3)	2.2(0.2)	1.8(0.2)	1.8(0.2)	15/15
2: CMA	2.9(1)	2.4(0.3)	2.3(0.3)	1.9(0.2)	1.9(0.2)	15/15
$\mathbf{f_{11}}$	143	763	1177	1467	1673	15/15
1: CMA	5.1(1)	**1.3**(0.2)$^{\star 3}$	**1.0**(0.1)$^{\star 3}$	**0.89**(0.1)$^{\star 3}$	**0.84**(0.1)$^{\star 3}$	15/15
2: CMA	6.5(2)	1.8(0.3)	1.3(0.2)	1.1(0.1)	1.0(0.1)	15/15
$\mathbf{f_{12}}$	108	371	461	1303	1494	15/15
1: CMA	6.1(3)	5.4(5)	5.9(4)	2.6(2)	2.6(2)	15/15
2: CMA	6.5(8)	8.1(6)	8.8(5)	3.9(2)	3.8(2)	15/15
$\mathbf{f_{13}}$	132	250	1310	1752	2255	15/15
1: CMA	3.1(2)	4.6(2)	1.2(0.3)	1.3(0.2)	1.2(0.2)	15/15
2: CMA	4.0(4)	5.6(3)	1.7(0.8)	1.7(0.6)	1.6(0.5)	15/15
$\mathbf{f_{14}}$	10	58	139	251	476	15/15
1: CMA	1.8(3)	3.3(0.6)	3.7(0.7)	4.0(0.8)	**3.0**(0.4)*	15/15
2: CMA	1.4(1)	**2.4**(0.6)*	3.4(0.9)	4.1(1)	3.5(0.4)	15/15
$\mathbf{f_{15}}$	511	19369	20073	20769	21359	14/15
1: CMA	1.0(0.5)	1.1(0.8)	1.1(0.8)	1.1(0.8)	1.1(0.8)	15/15
2: CMA	2.0(2)	1.1(0.7)	1.1(0.7)	1.1(0.7)	1.1(0.7)	15/15
$\mathbf{f_{16}}$	120	2662	10449	11644	12095	15/15
1: CMA	2.3(2)	1.7(0.8)	0.88(0.6)	0.83(0.5)	0.83(0.5)	15/15
2: CMA	2.3(1)	2.4(2)	0.94(0.7)	0.90(0.6)	0.90(0.6)	15/15
$\mathbf{f_{17}}$	5.2	899	3669	6351	7934	15/15
1: CMA	3.3(2)	1.1(1)	0.84(0.9)	0.82(0.5)	0.96(0.4)	15/15
2: CMA	3.1(2)	2.1(2)	1.3(0.8)	1.1(0.5)	1.0(0.4)	15/15
$\mathbf{f_{18}}$	103	3968	9280	10905	12469	15/15
1: CMA	0.89(0.7)	0.63(0.5)	0.69(0.6)	0.67(0.6)	0.71(0.5)	15/15
2: CMA	6.2(2)	1.5(1)	1.2(0.6)	1.1(0.7)	1.1(0.6)	15/15
$\mathbf{f_{19}}$	1	242	1.2e5	1.2e5	1.2e5	15/15
1: CMA	18(13)	259(255)	1.9(2)	1.9(2)	1.9(2)	15/15
2: CMA	17(12)	518(491)	2.4(2)	2.4(2)	2.3(2)	15/15
$\mathbf{f_{20}}$	16	38111	54470	54861	55313	14/15
1: CMA	2.4(2)	1.7(2)	1.3(1)	1.3(1)	1.4(1)	15/15
2: CMA	2.3(2)	1.6(2)	1.2(1)	1.2(1)	1.2(1)	15/15
$\mathbf{f_{21}}$	41	1674	1705	1729	1757	14/15
1: CMA	2.0(2)	22(18)	39(109)	39(107)	39(106)	13/15
2: CMA	3.6(4)	5.7(10)	5.9(11)	6.0(11)	6.1(11)	15/15
$\mathbf{f_{22}}$	71	938	1008	1040	1068	14/15
1: CMA	2.0(0.8)	250(369)	465(576)	452(558)	442(542)	7/15
2: CMA	6.3(14)	289(403)	346(523)	336(510)	329(362)	8/15
$\mathbf{f_{23}}$	3.0	14249	31654	33030	34256	15/15
1: CMA	2.5(2)	37(38)	22(32)	21(31)	20(29)	9/15
2: CMA	1.9(2)	52(71)	23(32)	22(31)	22(30)	9/15
$\mathbf{f_{24}}$	1622	6.4e6	9.6e6	1.3e7	1.3e7	3/15
1: CMA	1.3(1)	∞	∞	∞	∞ *1.0e6*	0/15
2: CMA	1.6(1)	∞	∞	∞	∞ *1.0e6*	0/15

20-D

Δf	1e+1	1e-1	1e-3	1e-5	1e-7	#succ
$\mathbf{f_1}$	43	43	43	43	43	15/15
1: CMA	5.8(0.8)	14(1)	23(1)	32(1)	41(2)	15/15
2: CMA	**3.8**(0.8)$^{\star 3}$	**10**(1)$^{\star 3}$	**15**(1)$^{\star 3}$	**21**(1)$^{\star 3}$	**27**(1)$^{\star 3}$	15/15
$\mathbf{f_2}$	385	387	390	391	393	15/15
1: CMA	22(4)	27(2)	29(2)	31(1)	32(1)	15/15
2: CMA	23(5)	28(2)	30(1)	31(1)	32(1)	15/15
$\mathbf{f_3}$	5066	7635	7643	7646	7651	15/15
1: CMA	7.1(3)	∞	∞	∞	∞ *2.7e6*	0/15
2: CMA	13(7)	4950(5826)	4945(5457)	4944(5724)	4941(5662)	1/15
$\mathbf{f_4}$	4722	7666	7700	7758	1.4e5	9/15
1: CMA	∞	∞	∞	∞	∞ *2.5e6*	0/15
2: CMA	∞	∞	∞	∞	∞ *2.5e6*	0/15
$\mathbf{f_5}$	41	41	41	41	41	15/15
1: CMA	4.6(1)	5.3(1)	5.3(1)	5.3(1)	5.3(1)	15/15
2: CMA	**3.1**(0.6)$^{\star 2}$	4.1(1)	4.1(1)	4.1(1)	4.1(1)	15/15
$\mathbf{f_6}$	1296	3413	5220	6728	8409	15/15
1: CMA	1.5(0.3)	1.0(0.1)	1.0(0.1)	1.1(0.1)	1.1(0.1)	15/15
2: CMA	**1.2**(0.3)*	1.0(0.2)	1.1(0.3)	1.2(0.3)	1.3(0.3)	15/15
$\mathbf{f_7}$	1351	9503	16524	16524	16969	15/15
1: CMA	1.6(1)	1.8(0.7)	1.1(0.4)	1.1(0.4)	1.1(0.4)	15/15
2: CMA	1.7(1)	1.5(0.6)	0.95(0.3)	0.95(0.3)	0.93(0.3)	15/15
$\mathbf{f_8}$	2039	4040	4219	4371	4484	15/15
1: CMA	3.1(0.6)	4.4(3)	4.5(3)	4.5(2)	4.5(2)	15/15
2: CMA	2.6(0.5)	3.6(2)	3.7(2)	3.6(2)	3.6(2)	15/15
$\mathbf{f_9}$	1716	3277	3455	3594	3727	15/15
1: CMA	3.4(0.9)	4.3(0.5)	4.4(0.5)	4.4(0.4)	4.4(0.4)	15/15
2: CMA	2.9(0.9)	4.9(3)	5.0(3)	4.9(3)	4.9(2)	15/15
$\mathbf{f_{10}}$	7413	10735	14920	17073	17476	15/15
1: CMA	1.1(0.2)	0.98(0.1)	0.76(0.0)$^{\downarrow 4}$	0.70(0.0)$^{\downarrow 4}$	0.71(0.0)$^{\downarrow 4}$	15/15
2: CMA	1.2(0.2)	1.0(0.1)	0.79(0.1)$^{\downarrow 4}$	0.71(0.0)$^{\downarrow 4}$	0.71(0.0)$^{\downarrow 4}$	15/15
$\mathbf{f_{11}}$	1002	6278	9762	12285	14831	15/15
1: CMA	**4.3**(0.5)*	**0.82**(0.1)$^{\star 2}$	**0.59**(0.0)$^{\star 2 \downarrow 4}$	**0.51**(0.0)$^{\star 2 \downarrow 4}$	**0.45**(0.0)$^{\star 2 \downarrow 4}$	15/15
2: CMA	4.8(0.5)	0.92(0.1)	0.66(0.0)$^{\downarrow 4}$	0.56(0.0)$^{\downarrow 4}$	0.49(0.0)$^{\downarrow 4}$	15/15
$\mathbf{f_{12}}$	1042	2740	4140	12407	13827	15/15
1: CMA	2.3(2)	3.2(2)	3.1(1)	1.3(0.5)	1.4(0.5)	15/15
2: CMA	**1.2**(1)*	2.6(2)	2.5(1)	1.1(0.4)	1.1(0.3)	15/15
$\mathbf{f_{13}}$	652	2751	18749	24455	30201	15/15
1: CMA	2.9(3)	4.4(2)	0.94(0.4)	1.1(0.5)	1.5(1)	15/15
2: CMA	3.6(3)	4.2(3)	1.3(0.4)	1.4(0.6)	1.8(0.9)	15/15
$\mathbf{f_{14}}$	75	304	932	1648	15661	15/15
1: CMA	3.3(1)	2.8(0.4)	2.9(0.3)	3.7(0.3)	0.65(0.0)$^{\downarrow 4}$	15/15
2: CMA	**2.0**(0.6)$^{\star 2}$	**1.9**(0.3)$^{\star 3}$	**2.3**(0.3)$^{\star 3}$	3.4(0.4)	0.66(0.0)$^{\downarrow 4}$	15/15
$\mathbf{f_{15}}$	30378	3.1e5	3.2e5	4.5e5	4.6e5	15/15
1: CMA	0.62(0.2)$^{\downarrow 2}$	0.65(0.3)	0.67(0.3)	0.49(0.2)$^{\downarrow 3}$	0.50(0.2)$^{\downarrow 2}$	15/15
2: CMA	0.92(0.6)	0.64(0.3)	0.65(0.3)	0.48(0.2)$^{\downarrow 2}$	0.49(0.2)$^{\downarrow 2}$	15/15
$\mathbf{f_{16}}$	1384	77015	1.9e5	2.0e5	2.2e5	15/15
1: CMA	2.2(3)	0.84(0.6)	0.76(0.6)	0.85(0.9)	0.79(0.9)	15/15
2: CMA	2.3(3)	0.90(0.6)	0.92(0.5)	1.3(1)	1.2(0.9)	15/15
$\mathbf{f_{17}}$	63	4005	30677	56288	80472	15/15
1: CMA	2.2(2)	**1.4**(2)*	0.70(0.3)	0.79(0.3)	0.82(0.2)$^{\downarrow}$	15/15
2: CMA	2.0(1)	4.0(2)	0.92(0.4)	0.96(0.4)	0.94(0.3)	15/15
$\mathbf{f_{18}}$	621	19561	67569	1.3e5	1.5e5	15/15
1: CMA	0.85(0.2)	0.78(0.7)	0.68(0.2)	0.77(0.4)	0.74(0.3)	15/15
2: CMA	0.95(0.5)	1.2(0.7)	0.79(0.3)	0.88(0.4)	0.91(0.3)	15/15
$\mathbf{f_{19}}$	1	3.4e5	6.2e6	6.7e6	6.7e6	15/15
1: CMA	135(44)	3.2(4)	1.0(1.0)	1.6(2)	2.1(2)	4/15
2: CMA	95(46)	2.5(4)	0.91(0.8)	1.3(1)	1.3(1)	6/15
$\mathbf{f_{20}}$	82	3.1e6	5.5e6	5.6e6	5.6e6	14/15
1: CMA	3.9(0.7)	1.1(0.7)	5.0(6)	5.0(5)	4.9(5)	2/15
2: CMA	**2.5**(0.8)$^{\star 2}$	0.87(0.4)	2.4(3)	3.2(4)	3.2(3)	3/15
$\mathbf{f_{21}}$	561	14103	14643	15567	17589	15/15
1: CMA	3.0(5)	95(126)	92(121)	86(115)	76(101)	7/15
2: CMA	25(22)	76(118)	73(116)	69(81)	61(72)	8/15
$\mathbf{f_{22}}$	467	23491	24948	26847	1.3e5	12/15
1: CMA	6.7(12)	∞	∞	∞	∞ *1.1e6*	0/15
2: CMA	187(29)	∞	∞	∞	∞ *1.1e6*	0/15
$\mathbf{f_{23}}$	3.2	67457	4.9e5	8.1e5	8.4e5	15/15
1: CMA	2.5(3)	255(296)	∞	∞	∞ *2.6e6*	0/15
2: CMA	2.9(3)	516(567)	∞	∞	∞ *2.4e6*	0/15
$\mathbf{f_{24}}$	1.3e6	5.2e7	5.2e7	5.2e7	5.2e7	3/15
1: CMA	19(22)	∞	∞	∞	∞ *4.0e6*	0/15
2: CMA	42(48)	∞	∞	∞	∞ *4.0e6*	0/15

Table 1: ERT **in number of function evaluations divided by the best** ERT **measured during BBOB-2009 given in the respective first row with the central 80% range divided by two in brackets for different Δf values. #succ is the number of trials that reached the final target $f_{\mathrm{opt}} + 10^{-8}$. 1:CMA is $\mathbf{CMA_{ma}}$ and 2:CMA is $\mathbf{CMA_{mah}}$. Bold entries are statistically significantly better compared to the other algorithm, with $p = 0.05$ or $p = 10^{-k}$ where $k \in \{2, 3, 4, \dots\}$ is the number following the $\star$ symbol, with Bonferroni correction of 48. A $\downarrow$ indicates the same tested against the best BBOB-2009.**